

# SCAN: a neural model of covert attention

Citation for published version (APA):

Postma, E. O. (1994). *SCAN: a neural model of covert attention*. [Doctoral Thesis, Maastricht University]. Rijksuniversiteit Limburg. <https://doi.org/10.26481/dis.19940922ep>

**Document status and date:**

Published: 01/01/1994

**DOI:**

[10.26481/dis.19940922ep](https://doi.org/10.26481/dis.19940922ep)

**Document Version:**

Publisher's PDF, also known as Version of record

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# Summary

While fixating the eyes on one location of the visual scene one is still able to identify objects at other locations. The process underlying this ability is called covert attention. The covert-attention process can be likened to a searchlight that selects part of a visual scene by illuminating it. The searchlight metaphor of covert attention is at the base of the SCAN model described in this thesis.

In Chapter 1, the problem statement of this study reads: *is it possible to construct a working model of covert attention as described by psychology which satisfies the constraints set by biology?*

A multidisciplinary approach is pursued in order to answer this question. The computer-science approach is central in that it emphasizes the formulation of a *working* model. In addition, psychological knowledge about the functional aspects of the attentional searchlight are combined with biological knowledge about its realisation in neural structures. Finally, models and insights from statistical physics are used in formulating and studying the behaviour of the model.

Scalability to problems of a realistic size is an important requirement for neural-network models. Minsky and Papert had already pointed out, in their book *Perceptrons*, that the performance of neural networks often remains limited to small-sized toy problems. Given that a working model has to be applicable to problems of a realistic size, it is important to take scalability into account when formulating the architecture. Biological knowledge about the neural realisation of brain function plays a pivotal role in this process. Characteristics of biological hardware may be formulated as constraints for scalable neural machines.

Chapter 2 studies the biological basis of visual perception and visual attention.

At the level of neurons and their interconnections three characteristics appear to be of importance for scalability: (1) the stochasticity of neurons, (2) the sparseness of connections, (3) and the profligacy of neurons. At the level of the visual system, two characteristics of vision are: (4) the hierarchical structure of the visual system, and (5) the invariant representation of visual objects in the identification part of the visual system. The characteristics (1) to (3) are used in the definition of the neural lattice in Chapter 3 and in defining the gating lattice in Chapter 4. The characteristics (4) and (5) are used in the SCAN model described in Chapter 5.

Chapter 3 studies the effect of connection density on neural-network performance. Because the scalability of a neural network decreases with increasing connectivity, it is important to keep the number and length of connections small. A model called the neural lattice is introduced and used to study the effect of connection density. The neural lattice is a parallel distributed neural network that detects an above-threshold input signal and is equivalent to the ferromagnetic Ising lattice from statistical mechanics. Although the neural lattice is not applied in the SCAN model, it is strongly related to the gating lattice, the building block of SCAN that is described in Chapter 4.

The structure and dynamics of the neural lattice obey the characteristics (1) to (3) of Chapter 2: the lattice consists of a large number of stochastic elements that communicate through nearest-neighbour connections. Varying the lattice dimensionality leads to three distinct models that differ in their connection density: the one-, two-, and three-dimensional neural lattices. The behaviour of the three lattices is compared with the behaviour of a lattice in which all  $N$  elements ( $N \gg 6$ ) are interconnected: the " $N$ -dimensional" neural lattice.

Through Monte-Carlo simulations the effect of lattice dimensionality on the robustness and flexibility of above-threshold detection are assessed. The robustness of a neural lattice is defined as: the degree to which the quality of detection is maintained with increasing noise in the input signal. The flexibility of a neural lattice is its independency from the initial state.

Simulation results show that the robustness increases with lattice dimensionality, whereas flexibility decreases with lattice dimensionality. This leads to the conclusion that optimal robustness in combination with optimal flexibility can be obtained for a two- or three-dimensional lattice.

Chapter 4 introduces the gating lattice, that serves as the building block for

SCAN in Chapter 5. The gating lattice is equivalent with the antiferromagnetic Ising lattice from statistical mechanics and is, therefore, related to the neural lattice. The conclusions about the neural lattice also apply, therefore, to the gating lattice. As a result the gating lattice may be defined as a two- or three- dimensional lattice. The gating lattice is defined as a two-dimensional lattice in order to be able to use existing knowledge of the equivalent antiferromagnetic Ising lattice.

The gating lattice is a parallel distributed switch with three switching directions. It is capable of selecting one of three patterns without affecting the structure of the pattern. The selection of a switching direction is based on the relative magnitudes of three control signals. The distributed competition in the gating lattice amplifies small differences among the control signals. As a result, the sublattice with the largest control signal opens. When the control signals are somehow derived from the pattern to be gated (e.g., its saliency), a mechanism is obtained in which a pattern (via its saliency) itself may determine whether it is gated or not.

A formulation of the gating lattice as an Ising model enables an exact computation of a lattice composed of nine elements. Monte-Carlo simulations are used for the study of larger lattices. Combining these results with knowledge of the critical behaviour of Ising models reveals that the optimal gating quality is achieved when the intrinsic-noise value is equal to a critical value. In contrast, the gating speed decreases when approaching the critical value. This leads to the conclusion that an intrinsic-noise value near the critical value yields the optimal combination of gating quality and gating speed.

Chapter 5 introduces the SCAN (Signal Channelling Attentional Network) model. Apart from the three first characteristics, SCAN incorporates the characteristics (4) and (5): it is a hierarchical network that yields a translation-invariant representation. SCAN achieves invariance by selecting contiguous subpatterns (object-patterns) from an input pattern. The gating network, a hierarchical structure of gating lattices, channels subpatterns towards a centralized neural-network classifier. The classifier used in SCAN is based on the ART network of Carpenter and Grossberg. The main feature of the ART network that is of relevance for our purposes is that it generates an expectation pattern when confronted with an input pattern. In SCAN, the expectation pattern is compared with all subpatterns of the input pattern. The degree to which the expectation and a subpattern match yields a control signal; the bet-

ter the match, the larger the control signal. Because a control signal controls the sublattice that gates the accompanying subpattern, the subpattern that has the best match with the expectation is channelled towards the classifier.

Simulations with random control signals show how the speed and quality of gating in SCAN behave as a function of the intrinsic-noise value and the number of levels in the gating tree. In order to enable the selection of patterns with an intrinsic two-dimensional structure (e.g., visual images), the gating lattice is defined as a fractal structure (see the cover). At each node in this structure, a gating lattice selects one of three subpatterns from a two-dimensional plane. When presented with a realistic image of  $3^{10}$  pixels and the associated control signals, SCAN is capable of channelling the target pattern with a reasonable overall gating quality. Perfect gating quality is obtained by varying the intrinsic noise in time.

It is concluded that the SCAN model constitutes a scalable realisation of the attentional searchlight. The method used in SCAN to obtain translation invariance differs radically from standard neural-network approaches. Generally, input images are decomposed in invariant features. Such a decomposition leads to the problem of how features belonging to a single object should be glued together, the so-called binding problem. In SCAN the binding problem does not arise because no decomposition takes place. Moreover, SCAN combines translation invariance with selection enabling and active scanning of visual images.

In Chapter 6, the structure and functioning of the SCAN model are evaluated by comparing it to the biological example. The plausibility of SCAN as a biological model is discussed. Two inconsistencies lead to suggestions for extending and improving the model.

Chapter 7 concludes that in SCAN a model is realised that is in agreement with the requirements set out in the problem statement by combining biology (the characteristics of a scalable architecture), psychology (the functionality of the attentional searchlight), statistical physics (the behaviour of a large number of locally interconnected stochastic elements), and computer science (the scalability of parallel distributed neurocomputers).

# Samenvatting

Wanneer de ogen gefixeerd worden op één lokatie in de visuele omgeving is het niettemin mogelijk objecten op andere lokaties te identificeren. Het proces dat hiervoor verantwoordelijk is wordt selectieve aandacht (*covert attention*) genoemd. Het visuele aandachtsproces kan tot op zekere hoogte vergeleken worden met een zoeklicht dat een deel van de omgeving selecteert door het te belichten. De notie van selectieve aandacht als zoeklicht ligt ten grondslag aan het in dit proefschrift beschreven SCAN-model.

In hoofdstuk 1 wordt de probleemstelling van deze studie geformuleerd. De probleemstelling luidt: *is het mogelijk een werkend model van visuele aandacht te formuleren dat in overeenstemming is met de functionaliteit van het aandachtszoeklicht zoals beschreven in de psychologie en dat voldoet aan de voorwaarden voor een succesvolle fysieke realisatie zoals beschreven in de biologie?*

Om deze vraag te beantwoorden is voor een multidisciplinaire aanpak gekozen. Centraal in deze aanpak staat de informatica-benadering waarin de nadruk valt op het formuleren van een *werkend* model. Daarnaast wordt psychologische kennis van de functionaliteit van het aandachtszoeklicht gecombineerd met biologische kennis over zijn realisatie in neurale structuren. Tenslotte worden modellen en inzichten uit de statistische fysica gebruikt voor de beschrijving en bestudering van het model.

Een belangrijk vereiste van een neuraal-netwerk model is dat het schaalbaar is naar problemen van een realistische omvang. Zoals reeds onderstreept door Minsky en Papert in hun boek *Perceptrons*, blijft het bereik van neurale netwerken vaak beperkt tot problemen van een kleine omvang. Aangezien een werkend model toepasbaar moet zijn op problemen van een realistische

omvang, is het van belang om bij het specificeren van de architectuur rekening te houden met de schaalbaarheid. Biologische kennis over de neurale realisatie van hersenfuncties speelt hierbij een belangrijke rol. Karakteristieke eigenschappen van de biologische hardware kunnen worden geformuleerd als voorwaarden voor schaalbare neurale machines.

In hoofdstuk 2 wordt de biologische basis van de visuele waarneming en de visuele aandacht bestudeerd. Op het niveau van neuronen en hun verbindingen blijken drie karakteristieke eigenschappen van belang te zijn voor schaalbaarheid: (1) de stochasticiteit van neuronen, (2) de zuinigheid van verbindingen, en (3) de veelheid van neuronen. Op het niveau van het visueel systeem komen twee karakteristieke eigenschappen naar voren die specifiek van toepassing zijn op de visuele waarneming: (4) de hiërarchische structuur van het visueel systeem en (5) de invariantie van object-representaties in het deel van het visueel systeem dat verantwoordelijk is voor object-identificatie. De eigenschappen (1) tot en met (3) worden gebruikt in de definitie van het neurale rooster in hoofdstuk 3 en in de definitie van het schakelrooster in hoofdstuk 4. De eigenschappen (4) en (5) worden gebruikt in het in hoofdstuk 5 gedefinieerde SCAN-model.

In hoofdstuk 3 wordt een studie verricht naar het effect van de mate van connectiviteit op het gedrag van een neurale netwerk. Aangezien de schaalbaarheid van een neurale netwerk afneemt met toenemende connectiviteit, is het van belang het de lengte van verbindingen en het aantal verbindingen klein te houden. Uitgangspunt voor de in dit hoofdstuk beschreven studie vormt het neurale rooster (*neural lattice*), een parallel gedistribueerd neurale netwerk dat een bovendrempelig invoersignaal detecteert. Het neurale rooster is gelijkwaardig aan het ferromagnetisch Ising rooster uit de statistische fysica. Hoewel het neurale rooster niet toegepast wordt in het SCAN-model, is het wel sterk verwant aan het in hoofdstuk 4 beschreven schakelrooster dat als bouwsteen van SCAN fungeert.

De structuur en dynamica van het neurale rooster voldoen aan de karakteristieken (1) tot en met (3) uit hoofdstuk 2: het rooster is opgebouwd uit een groot aantal stochastische elementen die door naaste-buur verbindingen communiceren. Variatie van de dimensionaliteit van het rooster leidt tot drie verschillende modellen die zich onderscheiden in hun mate van connectiviteit: het één-, twee-, en drie-dimensionale neurale rooster. Het gedrag van de drie roosters worden vergeleken met het gedrag van een rooster waarin alle  $N$  ele-

menten ( $N \gg 6$ ) met elkaar verbonden zijn: het " $N$ -dimensionale" rooster.

Door middel van Monte-Carlo simulaties wordt het effect van de dimensionaliteit van het rooster op de robuustheid en flexibiliteit van detectie bepaald. Onder de robuustheid van het neurale rooster wordt verstaan: de mate waarin de detectiekwaliteit behouden blijft bij toenemende ruis in het invoersignaal. De flexibiliteit van het neurale rooster is de mate waarin de detectiekwaliteit onafhankelijk is van de begintoestand van het rooster.

Uit de simulatieresultaten blijkt dat de robuustheid toeneemt met de dimensionaliteit van het rooster en de flexibiliteit afneemt met de dimensionaliteit van het rooster. Dit leidt tot de conclusie dat voor optimale robuustheid in combinatie met optimale flexibiliteit, een twee- of drie-dimensionaal rooster de voorkeur verdient boven het één- of  $N$ -dimensionaal rooster.

In hoofdstuk 4 wordt het schakelrooster (*gating lattice*) ingevoerd. Het schakelrooster is gelijkwaardig met het antiferromagnetisch Ising rooster en daarmee verwant aan het neurale rooster. De conclusies over het neurale rooster zijn derhalve tevens van toepassing op het schakelrooster. Als uitvloeisel van de conclusie van hoofdstuk 3, kan het schakelrooster gedefinieerd worden als een twee- of drie-dimensionaal rooster. Het schakelrooster wordt gedefinieerd als een twee-dimensionaal rooster zodat bestaande kennis van het gelijkwaardig antiferromagnetisch Ising rooster kan worden gebruikt.

Het schakelrooster vormt een parallel gedistribueerde drie-standen schakelaar die één van drie patronen selecteert waarbij de structuur van het patroon intact blijft. De keuze van de schakelrichting wordt bepaald door drie stuursignalen. De gedistribueerde competitie in het schakelrooster versterkt kleine verschillen tussen de stuursignalen zodat het rooster met het grootste stuursignaal zich opent. Wanneer de stuursignalen op een of andere wijze afgeleid zijn van de door te laten patronen (het stuursignaal drukt b.v. de opvallendheid van een patroon uit), ontstaat er een mechanisme waarbij het patroon zelf (via zijn opvallendheid) bepalend is voor het al of niet doorlaten ervan.

Een formulering van het schakelrooster als het Ising model maakt een exacte berekening mogelijk van een rooster bestaande uit negen elementen. Voor de bestudering van grotere roosters wordt gebruik gemaakt van Monte-Carlo simulaties. In combinatie met kennis over het kritische gedrag van het Ising model laten deze simulaties zien hoe de schakelkwaliteit en schakelsnelheid afhangen van de intrinsieke ruiswaarde. Het blijkt dat optimale schakel-



kwaliteit bereikt wordt wanneer de intrinsieke ruis een waarde heeft die gelijk is aan de kritische waarde. De schakelsnelheid neemt daarentegen af naarmate de intrinsieke ruis de kritische waarde dichtert nadert. Dit leidt tot de conclusie dat door een waarde van de intrinsieke ruis nabij de kritische waarde een optimale combinatie van schakelkwaliteit en schakelsnelheid verkregen wordt. Het schakelrooster vormt de bouwsteen van het in hoofdstuk 5 beschreven SCAN-model.

In hoofdstuk 5 wordt het SCAN (Signal Channelling Attentional Network) model geïntroduceerd. In het SCAN-model zijn, naast de eerste drie karakteristieke eigenschappen, de eigenschappen (4) en (5) verwerkt: het vormt een hiërarchisch netwerk dat leidt tot een translatie invariante representatie. SCAN bereikt invariantie door aaneengesloten deelpatronen (objectpatronen) te selecteren uit een invoerpatroon. Dit geschiedt door middel van een schakelnetwerk (*gating network*), een hiërarchische structuur van schakelroosters, die deelpatronen in baan brengt naar een gecentraliseerde neurale-netwerk patroonclassificator. De in SCAN gebruikte classificator is gebaseerd op het ART-netwerk van Carpenter en Grossberg. De voor deze toepassing voorname eigenschap van het ART-netwerk is dat het een verwachting genereert voor een voor classificatie aangeboden invoerpatroon. In SCAN wordt het door de classificator gegenereerde verwachtingspatroon vergeleken met alle deelpatronen in het invoerpatroon. De mate van overeenkomst van een deelpatroon met het verwachtingspatroon levert de waarde van het stuursignaal: hoe beter de overeenkomst, des te hoger het stuursignaal. Aangezien dit stuursignaal gekoppeld is aan het subrooster dat het betreffend deelpatroon als invoer heeft, wordt het deelpatroon dat het meest met de verwachting overeenkomt het invoerpatroon van de classificator.

Simulaties met willekeurige stuursignalen laten zien hoe SCAN zich gedraagt, afhankelijk van de waarde van de intrinsieke ruis, en hoe de schakelkwaliteit en schakelsnelheid afhangen van het aantal lagen in het schakelnetwerk. Om selectie uit invoerpatronen met een intrinsiek twee-dimensionale structuur (b.v. een foto) mogelijk te maken, wordt het schakelnetwerk gedefinieerd als een fractale structuur (zie omslag). Op ieder knooppunt van deze structuur selecteert een schakelrooster één van drie deelpatronen uit het twee-dimensionale vlak. Gegeven een realistisch visueel beeld bestaande uit  $3^{10}$  beeldelementen en de bijbehorende stuursignalen, is SCAN in staat om het deelpatroon met een redelijke schakelkwaliteit in baan te brengen naar de classificator. Een perfecte schakelkwaliteit wordt bereikt door de intrinsieke ruis in de tijd te

variëren.

De conclusie van hoofdstuk 5 is dat het SCAN-model een schaalbare realisatie is van het aandachtszoeklicht. De in SCAN gehanteerde methode om translatie-invariante patroonherkenning te bereiken wijkt radicaal af van de gangbare neurale-netwerk methoden. Doorgaans wordt namelijk uitgegaan van een ontbinding van het beeld in invariante kenmerken. Een dergelijke ontbinding leidt tot het bindingsprobleem (*binding problem*), oftewel, het probleem hoe de kenmerken behorend bij een object verbonden worden tot een invariante objectrepresentatie. In SCAN doet het bindingsprobleem zich niet voor, omdat er geen ontbinding plaatsvindt. Bovendien combineert SCAN translatie-invariantie met selectie waardoor een actieve *scanning* van visuele beelden mogelijk wordt.

In hoofdstuk 6 worden de structuur en het gedrag van het SCAN-model geëvalueerd door een vergelijking met het biologische voorbeeld. De aanvaardbaarheid van SCAN als biologisch model wordt besproken. Twee inconsistenties leiden tot suggesties voor de wijze waarop het SCAN-model uitgebreid en verbeterd kan worden.

In hoofdstuk 7 wordt geconcludeerd dat in SCAN een model verwezenlijkt is dat door de combinatie van biologie (de karakteristieke kenmerken van een schaalbare architectuur), psychologie (de functionaliteit van het aandachtszoeklicht), statistische fysica (het gedrag van een groot aantal stochastische lokaal communicerende elementen), en informatica (de schaalbaarheid van parallel gedistribueerde neurocomputers) in overeenstemming is met de in de probleemstelling genoemde eisen.